Symmetry, Integrability and Geometry: Methods and Applications SIGMA **11** (2015), 077, 10 pages

Moments and Legendre–Fourier Series for Measures Supported on Curves*

Jean B. LASSERRE

LAAS-CNRS and Institute of Mathematics, University of Toulouse,
7 Avenue du Colonel Roche, BP 54 200, 31031 Toulouse Cédex 4, France
E-mail: *lasserre@laas.fr*
URL: http://homepages.laas.fr/lasserre/

Received August 28, 2015, in final form September 26, 2015; Published online September 29, 2015
http://dx.doi.org/10.3842/SIGMA.2015.077

Abstract. Some important problems (e.g., in optimal transport and optimal control) have a relaxed (or weak) formulation in a space of appropriate measures which is much easier to solve. However, an optimal solution μ of the latter solves the former if and only if the measure μ is supported on a "trajectory" $\{(t, x(t))\colon t \in [0, T]\}$ for some measurable function $x(t)$. We provide necessary and sufficient conditions on moments (γ_{ij}) of a measure $d\mu(x, t)$ on $[0, 1]^2$ to ensure that μ is supported on a trajectory $\{(t, x(t))\colon t \in [0, 1]\}$. Those conditions are stated in terms of Legendre–Fourier coefficients $\mathbf{f}_j = (\mathbf{f}_j(i))$ associated with some functions $f_j\colon [0, 1] \to \mathbb{R}$, $j = 1, \ldots$, where each $\mathbf{f}_j$ is obtained from the moments γ_{ji}, $i = 0, 1, \ldots$, of μ.

Key words: moment problem; Legendre polynomials; Legendre–Fourier series

2010 Mathematics Subject Classification: 42C05; 42C10; 42A16; 44A60

1 Introduction

This paper is in the line of research concerned with the following issue: *which type and how much of information on the support of a measure can be extracted from its moments* (a research issue outlined in a *Problem session* at the 2013 Oberwolfach meeting on *Structured Function Systems and Applications* [2]). In particular, a highly desirable result is to obtain necessary and/or sufficient conditions on moments of a given measure to ensure that its support has certain geometric properties. For instance there is a vast literature on the old and classical L-moment problem, which asks for moment conditions to ensure that the underlying measure μ is absolutely continuous with respect to some reference measure ν, and with a density in $L_\infty(\nu)$. See, for instance, [3, 10, 11], more recently [7], and the many references therein.

Here we are interested in a problem that is somehow "orthogonal" to the L-moment problem. Namely, we consider the following generic problem: Let $d\mu(x, t)$ be a probability measure on $[0, 1] \times [0, 1]$. Provide necessary and/or sufficient conditions on the moments of μ to ensure that μ is singular with respect to the Lebesgue measure $d(x, t)$ on $[0, 1]^2$. In fact, and more precisely, suppose that:

- one knows all moments $\gamma_i(j) = \int x^i t^j \, d\mu(x, t)$, $i, j = 0, 1, \ldots$, of the measure μ, and
- the marginal of μ with respect to the "t" variable is the Lebesgue measure dt on $[0, 1]$.

Then provide necessary and/or sufficient conditions on the moments $(\gamma_i(j))$ of μ to ensure that μ is supported on a trajectory $\{(t, x(t))\colon t \in [0, 1]\} \subset [0, 1]^2$, for some measurable function $x\colon [0, 1] \to [0, 1]$.

*This paper is a contribution to the Special Issue on Orthogonal Polynomials, Special Functions and Applications. The full collection is available at http://www.emis.de/journals/SIGMA/OPSFA2015.html

In contrast to the L-moment problem, and to the best of our knowledge, the above problem stated in this form has not received a lot of attention in the past even though it is crucial in some important applications (two of them having motivated our interest).

Motivation. In addition of being of independent interest, this investigation is motivated by at least two important applications:

– *The mass transfer* (*or optimal transport*) *problem.* In the weak (or relaxed) Monge–Kantorovich formulation of the mass transport problem originally stated by Monge, one searches for a measure $d\mu(x,t)$ with prescribed marginals ν_x and ν_t, and which minimizes some cost functional $\int c(x,t)\,d\mu(x,t)$. However in the original Monge formulation, ultimately one would like to obtain an optimal solution μ^* of the form $d\mu^*(x,t) = \delta_{x(t)}\,d\nu_t(t)$ for some measurable function $t \mapsto x(t)$ (the transportation plan) and a crucial issue is to provide conditions for this to happen[1]. For more details the interested reader is referred, e.g., to [13, pp. 1–5] and [9]. There exist some characterizations of the support of an optimal measure for the weak formulation. For instance, *c-cyclical monotonicity* relates optimality with the support of solutions, and more recently [1] have shown in the (more general) context of the generalized moment problem that under some weak conditions the support of optimal solutions is *finitely minimal* / *c-monotone*. (As defined in [1] a set Γ is called finitely minimal / c-monotone if each finite measure α concentrated on finitely many atoms of Γ is cost minimizing among its competitors; in the optimal transport context, a competitor of α is any finite measure α' with same marginals as α.) For more details the interested reader is referred to [1] and the references therein. But such a characterization does not say when this support is a trajectory.

– *Deterministic optimal control.* Using the concept of *occupation measures*, a *weak formulation* of deterministic optimal control problems replaces the original control problem with an infinite-dimensional optimization problem $\mathcal{P}$ on a space of appropriate (occupation) measures on a Borel space $\mathscr{X} \times \mathscr{U} \times [0,1]$ with $\mathscr{X} \subset \mathbb{R}^n$, $\mathscr{U} \subset \mathbb{R}^m$. For more details the interested reader is referred, e.g., to [8, 14], and the many references therein. An important issue is to provide conditions on the problem data under which the optimal value of the relaxed problem $\mathcal{P}$ is the same as that of the original problem; see, e.g., [14]. Again this is the case if some optimal solution μ^* (or every element of a minimizing sequence) of the relaxed problem is such that every marginal μ^*_j of μ^* with respect to (x_j,t), $j = 1,\dots,n$, and every marginal μ^*_ℓ of μ^* with respect to (u_ℓ,t), $\ell = 1,\dots,m$, is supported on a trajectory $\{(t,x_j(t))\colon t \in [0,1]\}$ and on a trajectory $\{(t,u_\ell(t))\colon t \in [0,1]\}$ for some measurable functions $t \mapsto x_j(t)$ and $t \mapsto u_\ell(t)$ on $[0,1]$.

Contribution. Of course there is a particular case where one may conclude that μ is singular with respect to the Lebesgue measure on $[0,1]^2$. If there is a polynomial $p \in \mathbb{R}[x,t]$ of degree say d, such that its vector of coefficients $\mathbf{p}$ is in the kernel of the moment matrix $\mathbf{M}_s$ (where $\mathbf{M}_s[(i,j),(k,\ell)] = \gamma_{i+k,j+\ell}$, $i+j, k+\ell \le s$, with $d \le s$), then μ is supported on the variety $\{(x,t) \in [0,1]^2\colon p(x,t) = 0\}$ and therefore is singular with respect to the Lebesgue measure on $[0,1]^2$. But it may happen that $p(x,t) = p(y,t) = 0$ for some t and some $x \neq y$ and so even in this case additional conditions are needed to ensure existence of a trajectory $\{(t,x(t))\colon t \in [0,1]\}$.

We provide a set of explicit necessary and sufficient conditions on the moments $\boldsymbol{\gamma}_i = (\gamma_i(j))$ which state that for every fixed i, the moments $\gamma_i(j)$, $j = 0,1,\dots$, are limits of certain i-powers of the moments $\boldsymbol{\gamma}_1$.

More precisely, an explicit linear transformation $\Delta\boldsymbol{\gamma}_1$ of the infinite vector $\boldsymbol{\gamma}_1$ is the vector of (shifted) Legendre–Fourier coefficients associated with the function $t \mapsto x(t)$. Then the conditions state that for each fixed $i = 2,3,\dots$, the vector $\Delta\boldsymbol{\gamma}_i$ should be the vector of (shifted) Legendre–Fourier coefficients associated with the function $t \mapsto x(t)^i$, which in turn are expressible in terms of limits of "i-powers" of coefficients of $\Delta\boldsymbol{\gamma}_1$.

At last but not least, it should be noted that all results of this paper are easily transposed to the multi-dimensional case of a measure $d\mu(\mathbf{x},t)$ on $[0,1]^n \times [0,1]$ and supported on a trajectory

[1] Here δ_z denote the Dirac measure at the point z.

$\{(t,\mathbf{x}(t))\colon t\in[0,1]\}\subset[0,1]^{n+1}$ for some measurable mapping $\mathbf{x}\colon[0,1]\to[0,1]^n$. Indeed by proceeding *coordinate-wise* for each function $t\mapsto x_i(t)$, $i=1,\ldots,n$, one is reduced to the case $[0,1]^2$ investigated here.

2 Notation, definitions and preliminary results

2.1 Notation and definitions

Given a Borel probability measure μ on $[0,1]^2$, define

$$\gamma_i(j)=\int_{[0,1]^2} t^j x^i\,d\mu(x,t),\qquad i,j=0,1,\ldots, \tag{2.1}$$

and for every fixed $i\in\mathbb{N}$, denote by $\boldsymbol{\gamma}_i$ the vector of moments $(\gamma_i(j))$, $j=0,1,\ldots$.

Let $(\mathscr{L}_j)$, $j=0,1,\ldots$, be the family of orthonormal polynomials with respect to the Lebesgue measure on $[0,1]$. They can be deduced from the Legendre polynomials[2] via the change of variable $t'=(2t-1)$; the $(\mathscr{L}_j)$ are called the *shifted Legendre polynomials*. The polynomials $(\mathscr{L}_j)$ can be also computed exactly from the moments $\boldsymbol{\gamma}_0$ of the Lebesgue measure dt on $[0,1]$ by computing some determinants of modified Hankel moment matrices. For instance, $\mathscr{L}_0=1$, and

$$\mathscr{L}_1(t)=a\det\begin{pmatrix}1 & 1/2\\ 1 & t\end{pmatrix}=a(t-1/2)\qquad\text{with}\quad a^2(1/3-1/2+1/4)=1,$$

i.e., $a=\sqrt{12}$ and $\mathscr{L}_1(t)=2\sqrt{3}t-\sqrt{3}$, and

$$\mathscr{L}_2(t)=b\det\begin{pmatrix}1 & 1/2 & 1/3\\ 1/2 & 1/3 & 1/4\\ 1 & t & t^2\end{pmatrix}=b\big[t^2/12-t/12+1/72\big],$$

with $b>0$ such that $\int_0^1\mathscr{L}_2(t)^2dt=1$. See, e.g., [4] and [6].

The Lebesgue space $L_2([0,1]):=\{f\colon[0,1]\to\mathbb{R},\ \int_0^1 f^2\,dx<\infty\}$, equipped with the scalar product $\langle f,g\rangle=\int_0^1 fg\,dx$ and the associated norm $\|\cdot\|$, is a Hilbert space and the polynomials are dense in $L_2([0,1])$. In particular the family $(\mathscr{L}_j)$, $j=0,1,\ldots$, form an orthonormal basis of $L_2([0,1])$. Let ℓ^2 denotes the space of square-summable sequences with norm also denoted by $\|\cdot\|$.

Finally, let $\|f\|_\infty:=\operatorname{ess}\sup\limits_{x\in[0,1]}|f(x)|$, and similarly, for $p\in\mathbb{R}[x]$ let $\|p\|_\infty:=\sup\limits_{x\in[0,1]}|p(x)|$. Then $L_\infty([0,1]):=\{f\colon\|f\|_\infty<\infty\}$.

2.2 Some preliminary results

We next state some useful auxiliary results, some of them being standard in Real Analysis.

Proposition 2.1. *Let $t\mapsto f(t)$ be an element of $L_2([0,1])$ and define $\mathbf{f}=(\mathbf{f}(j))$ by*

$$\mathbf{f}(j):=\int_0^1\mathscr{L}_j(t)f(t)\,dt,\qquad j=0,1,\ldots.$$

Then one has

$$\sum_{j=0}^\infty\mathbf{f}(j)\mathscr{L}_j\left(:=\lim_{n\to\infty}\sum_{j=0}^n\mathbf{f}(j)\mathscr{L}_j\right)=f\qquad\text{in}\quad L_2([0,1]), \tag{2.2}$$

and this decomposition is unique. Moreover $\mathbf{f}\in\ell^2$ and $\|f\|=\|\mathbf{f}\|$.

[2] The Legendre polynomials are orthonormal w.r.t. to the Lebesgue measure on $[-1,1]$.

When the interval is $[-1,1]$ (instead of $[0,1]$ here) (2.2) is called the Legendre (or Legendre–Fourier) series expansion of the function f and $\mathbf{f}=(\mathbf{f}(j))$ is called the vector of Legendre–Fourier coefficients.

The notation f^k stands for the function $t\mapsto f(t)^k$, $k\in\mathbb{N}$. If $f^k\in L_2([0,1])$ we denote by $\mathbf{f}_k=(\mathbf{f}_k(j))\in\ell^2$ its (shifted) Legendre–Fourier coefficients so that $\|f^k\|=\|\mathbf{f}_k\|$ (where again the latter norm is that of ℓ^2). Notice that we also have:

Proposition 2.2. *Let $f,f^k,g\in L_2([0,1])$ with $k\in\mathbb{N}$. Then $g=f^k$ if and only if $\hat{\boldsymbol{g}}=\hat{\boldsymbol{f}}_k$.*

This follows from the uniqueness of the decomposition in the basis $(\mathscr{L}_j)$.

We also have the following helpful results:

Lemma 2.3. *Let $f\in L_2([0,1])$ with Legendre–Fourier coefficients $\mathbf{f}=(\mathbf{f}(j))$, and let $(p_n)\subset\mathbb{R}[t]$ be a sequence of polynomials such that $\|p_n-f\|\to 0$ as $n\to\infty$.*

If $\hat{\boldsymbol{p}}_n=(\hat{\boldsymbol{p}}_n(j))$ denotes the (shifted) Legendre–Fourier coefficients of p_n, for all $n=1,2,\ldots$, then $\|\hat{\boldsymbol{p}}_n-\mathbf{f}\|\to 0$ in ℓ^2 as $n\to\infty$.

Proof. As $\|p_n-f\|^2=\int_0^1(p_n-f)^2\,dt$ and with $d_n=\deg(p_n)$,

$$\begin{aligned}\int_0^1(p_n-f)^2\,dt &= \int_0^1\left(\sum_{j=0}^{d_n}\hat{\boldsymbol{p}}_n(j)\mathscr{L}_j-f\right)^2 dt\\ &=\sum_{j=0}^{d_n}(\hat{\boldsymbol{p}}_n(j))^2\underbrace{\int_0^1\mathscr{L}_j^2\,dt}_{=1}+2\sum_{k<j}\hat{\boldsymbol{p}}_n(j)\hat{\boldsymbol{p}}_n(k)\underbrace{\int_0^1\mathscr{L}_j\mathscr{L}_k\,dt}_{=0}\\ &\quad-2\sum_{j=0}^{d_n}\hat{\boldsymbol{p}}_n(j)\underbrace{\int_0^1\mathscr{L}_jf\,dt}_{\mathbf{f}(j)}+\underbrace{\|f\|^2}_{=\|\mathbf{f}\|=\sum_j\mathbf{f}(j)^2}\\ &\geq\sum_{j=0}^{d_n}\hat{\boldsymbol{p}}_n(j)^2-2\hat{\boldsymbol{p}}_n(j)\mathbf{f}(j)+\mathbf{f}(j)^2=\sum_{j=0}^{d_n}(\hat{\boldsymbol{p}}_n(j)-\mathbf{f}(j))^2,\end{aligned}$$

and the result follows because $\|p_n-f\|\to 0$ as $n\to\infty$. ∎

Lemma 2.4. *Let $f\in L_\infty([0,1])$ (hence $f^k\in L_2([0,1])$ for every $k\in\mathbb{N}$). If a sequence $(p_n)\subset\mathbb{R}[x]$ is such that $\sup_n\|p_n\|_\infty<\infty$ and $\|p_n-f\|\to 0$ as $n\to\infty$, then for every $k\in\mathbb{N}$,*

$$\lim_{n\to\infty}\left\|p_n^k-f^k\right\|=0.$$

In addition, if $\hat{\boldsymbol{p}}_n^k$ denotes the (shifted) Legendre–Fourier coefficients of p_n^k, $n=1,\ldots$, then $\|\hat{\boldsymbol{p}}_n^k-\mathbf{f}_k\|\to 0$ in ℓ^2 as $n\to\infty$.

Proof. Let $M>\max[\|f\|_\infty,\sup_n\|p_n\|_\infty]$ and fix $k\in\mathbb{N}$,

$$\begin{aligned}\|p_n^k-f^k\|^2 &= \int_0^1\left(p_n^k-f^k\right)^2dx=\int_0^1(p_n-f)^2\left(\sum_{\ell=0}^{k-1}p_n^{k-1-\ell}f^\ell\right)^2dx\\ &\leq\int_0^1(p_n-f)^2\left(\sum_{\ell=1}^{k}\left|p_n^{k-\ell}f^{\ell-1}\right|\right)^2dx\leq\left(kM^{k-1}\right)^2\int_0^1(p_n-f)^2\,dx\\ &=\left(kM^{k-1}\right)^2\|p_n-f\|^2\quad(\to 0\text{ as }n\to\infty.)\end{aligned}$$

Then the last statement follows from Lemma 2.3. ∎

Definition 2.5. Let $f \in L_2([0,1])$ with Legendre–Fourier coefficients $\mathbf{f}$. For every $k, n \in \mathbb{N}$, define the polynomial $f_n^{(k)} \in \mathbb{R}[x]$ and the vector $\mathbf{f}_n^{(k)} \in \mathbb{R}^{kn+1}$ by

$$t \mapsto f_n^{(k)}(t) := \left(\sum_{j=0}^{n} \mathbf{f}(j)\mathscr{L}_j(t)\right)^k = \sum_{j=0}^{nk} \mathbf{f}_n^{(k)}(j)\mathscr{L}_j(t).$$

Observe that each entry $\mathbf{f}_n^{(k)}(j)$, $j = 0, \ldots, nk$, is a degree-k form of the first $n+1$ Legendre–Fourier coefficients of $\hat{\boldsymbol{f}}$. Completing with zeros, consider $\mathbf{f}_n^{(k)}$ to be an element of ℓ^2 and if $\mathbf{f}_n^{(k)}$ converges in ℓ^2 as $n \to \infty$, call $\mathbf{f}^{(k)} \in \ell^2$ its limit.

The limit $\mathbf{f}^{(k)}$ can also be denoted $\mathbf{f} \star \cdots \star \mathbf{f}$, the limit of the k times "$\star$-product" in ℓ^2 of the vector $\mathbf{f} \in \ell^2$ by itself. Equivalently one may write $\mathbf{f}^{(k)} = \mathbf{f}^{(k-1)} \star \mathbf{f}$ since $f_n^{(k)}(t) = f_n^{(k-1)}(t) f_n^{(1)}(t)$ for all $t \in [0,1]$, and $\mathbf{f}_n^{(k-1)} \to \mathbf{f}^{k-1}$ as $n \to \infty$, as well as $\mathbf{f}_n^{(1)} \to \mathbf{f}$.

Lemma 2.6. *Let $f \in L_\infty([0,1])$, hence $f^k \in L_2([0,1)$ with (shifted) Legendre–Fourier coefficients $\mathbf{f}_k \in \ell^2$ for every $k \in \mathbb{N}$, and assume that*

$$\sup_n \|f_n^{(1)}\|_\infty \quad \left(= \sup_n \|\sum_{j=0}^{n} \mathbf{f}(j)\mathscr{L}_j\|_\infty\right) < \infty.$$

Then $\mathbf{f}_k = \mathbf{f}^{(k)} = \mathbf{f} \star \cdots \star \mathbf{f}$ (k times) for every $k = 1, 2, \ldots$, meaning that for every fixed $k \in \mathbb{N}$,

$$\lim_{n\to\infty} \left\| \left(\sum_{j=0}^{n} \mathbf{f}(j)\mathscr{L}_j\right)^k - f^k \right\| \quad \left(= \lim_{n\to\infty} \left\| \sum_{j=0}^{kn} \mathbf{f}_n^{(k)}(j)\mathscr{L}_j - f^k \right\|\right)$$

$$= \lim_{n\to\infty} \left\| \sum_{j=0}^{n} \mathbf{f}_k(j)\mathscr{L}_j - f^k \right\| = 0.$$

Equivalently, $\mathbf{f}_k = \mathbf{f}_{k-1} \star \mathbf{f}$ for every $k = 2, 3, \ldots$.

Proof. The result is a direct consequence of Lemmas 2.3 and 2.4 with $p_n = f_n^{(1)}$ (and the definition of the limit "$\star$-product" in Definition 2.5). ∎

3 Main result

Assume that we are given all moments of a nonnegative measure $d\mu(x,t)$ on a box $[a,b] \times [c,d] \subset \mathbb{R}^2$. After a re-scaling of its moments we may and will assume that μ is a probability measure supported on $[0,1]^2$ with associated moments

$$\gamma_i(j) = \int_{[0,1]^2} x^i t^j \, d\mu(x,t), \qquad i,j = 0,1,\ldots.$$

We further assume that the marginal measure μ_t with respect to the variable t, is the Lebesgue measure on $[0,1]$, that is, $\gamma_0(j) = 1/(j+1)$, $j = 0,1,\ldots$.

A standard disintegration of the measure μ yields

$$\gamma_i(j) = \int_{[0,1]^2} t^j x^i \, d\mu(x,t) = \int_0^1 t^j \underbrace{\left(\int_{[0,1]} x^i \psi(dx|t)\right)}_{=: f_i(t)} dt, \qquad i,j = 0,1,\ldots, \tag{3.1}$$

where the stochastic kernel $\psi(\cdot|\,t)$ is the conditional probability on $[0,1]$ given $t\in[0,1]$. Observe that the measurable function f_i in (3.1) is nonnegative and uniformly bounded by 1 because $|x^i|\leq 1$ on $[0,1]$ for every i, and so $f_i\in L_\infty([0,1])$ for every $i=1,\ldots$.

The vector $\boldsymbol{\gamma}_i=(\gamma_i(j))$, $j=0,1,\ldots$, is the vector of moments of the measure $d\mu_i(t)=f_i(t)dt$ on $[0,1]$, for every $i=1,2,\ldots$. The (shifted) Legendre–Fourier vector of coefficients $\hat{\boldsymbol{q}}_i$ of f_i are obtained easily from the (infinite) vector $\boldsymbol{\gamma}_i$ via a triangular linear system. Indeed write

$$\mathscr{L}_j(t)=\sum_{k=0}^{j}\Delta_{jk}t^k,\qquad \forall\, t\in[0,1],\quad j=0,1,\ldots,$$

where $\Delta_{jj}>0$, or in compact matrix form

$$\begin{bmatrix}\mathscr{L}_0\\ \mathscr{L}_1\\ \cdot\\ \mathscr{L}_n\\ \cdot\end{bmatrix}=\boldsymbol{\Delta}\begin{bmatrix}1\\ t\\ \cdot\\ t^n\\ \cdot\end{bmatrix},\tag{3.2}$$

for some infinite lower triangular matrix Δ with all diagonal elements being strictly positive. Therefore

$$\boldsymbol{\Delta}\boldsymbol{\gamma}_i=\boldsymbol{\Delta}\int_0^1\begin{bmatrix}1\\ t\\ \cdot\\ t^n\\ \cdot\end{bmatrix}f_i(t)\,dt=\int_0^1\begin{bmatrix}\mathscr{L}_0\\ \mathscr{L}_1\\ \cdot\\ \mathscr{L}_n\\ \cdot\end{bmatrix}f_i(t)\,dt=\hat{\boldsymbol{q}}_i.$$

Suppose that the measure μ is supported on a trajectory $\{(t,x(t))\colon t\in[0,1]\}\subset[0,1]^2$ for some measurable (density) function $x\colon[0,1]\to[0,1]$. The measurable function $t\mapsto x(t)$ is an element of $L_\infty([0,1])$ because $\|x\|_\infty\leq 1$. Then by Proposition 2.1,

$$x=\sum_{j=0}^{\infty}\hat{\boldsymbol{x}}(j)\mathscr{L}_j\qquad\text{in } L_2([0,1]),\tag{3.3}$$

where $\hat{\boldsymbol{x}}=(\hat{\boldsymbol{x}}(j))\in\ell^2$ is its vector of (shifted) Legendre–Fourier coefficients (with $\|\hat{\boldsymbol{x}}\|=\|x\|$). Similarly, for every $k=2,3,\ldots$, the function $t\mapsto x(t)^k$ is in $L_\infty([0,1])$ and

$$x^k=\sum_{j=0}^{\infty}\hat{\boldsymbol{x}}_k(j)\mathscr{L}_j\qquad\text{in } L_2([0,1]),$$

with vector of (shifted) Legendre–Fourier coefficients $\hat{\boldsymbol{x}}_k\in\ell^2$ such that $\|x^k\|=\|\hat{\boldsymbol{x}}_k\|$.

We also recall the notation $\hat{\boldsymbol{x}}_n^{(k)}\in\mathbb{R}^{kn+1}$ for the vector of coefficients in the basis $(\mathscr{L}_j)$ of the polynomial $t\mapsto\left(\sum_{j=0}^{n}\hat{\boldsymbol{x}}(j)\mathscr{L}_j(t)\right)^k$, and when considered as an element of ℓ^2 (by completing with zeros) denote by $\hat{\boldsymbol{x}}^{(k)}\in\ell^2$ its limit when it exists.

Theorem 3.1. *Let μ be a Borel probability measure on $[0,1]^2$ and let $\boldsymbol{\gamma}_i(j)$, $i,j=0,1,\ldots$, be the moments of μ in (2.1).*

(a) *If μ is supported on a trajectory $\{(t,x(t))\colon t\in[0,1]\}$ for some nonnegative measurable function $t\mapsto x(t)$ on $[0,1]$ and if* $\sup_n \left\|\sum_{j=0}^n \hat{\boldsymbol{x}}(j)\mathscr{L}_j\right\|_\infty < \infty$, *then*

$$\hat{\boldsymbol{x}}_i = \hat{\boldsymbol{x}}^{(i)} = \underbrace{\hat{\boldsymbol{x}}\star\cdots\star\hat{\boldsymbol{x}}}_{i\ \text{times}} = \hat{\boldsymbol{x}}_{i-1}\star\hat{\boldsymbol{x}}, \qquad \forall\, i=2,3,\ldots, \tag{3.4}$$

Equivalently

$$\boldsymbol{\Delta\gamma}_i = (\boldsymbol{\Delta\gamma}_1)^{(i)} = \left(\boldsymbol{\Delta\gamma}_{i-1}\star\boldsymbol{\Delta\gamma}_1\right)^{(i)}, \qquad \forall\, i=2,3,\ldots, \tag{3.5}$$

where $\boldsymbol{\Delta}$ is the non singular triangular matrix defined in (3.2).

(b) *Conversely, if (3.5) holds then μ is supported on a trajectory $\{(t,x(t))\colon t\in[0,1]\}$ for some measurable function $t\mapsto x(t)$ on $[0,1]$, and (3.4) also holds.*

Proof. The (a) part. As μ is supported on $[0,1]^2$ one has $\|x\|_\infty\le 1$ and so the function $t\mapsto x(t)^i$ is in $L_2([0,1])$ for every $i=1,2,\ldots$. So let $t\mapsto x(t)$ be written as in (3.3). Consider the function $t\mapsto x(t)^i$, for every fixed $i\in\mathbb{N}$, so that

$$x^i = \sum_{j=0}^\infty \hat{\boldsymbol{x}}_i(j)\mathscr{L}_j \qquad \text{in } L_2([0,1]),$$

where the (shifted) Legendre–Fourier vector of coefficients $\hat{\boldsymbol{x}}_i$ is obtained by $\hat{\boldsymbol{x}}_i = \boldsymbol{\Delta}^{-1}\boldsymbol{\gamma}_i$. But by Lemma 2.6, we also have

$$x^i = \left(\sum_{j=0}^\infty \hat{\boldsymbol{x}}(j)\mathscr{L}_j\right)^i \left(=\lim_{n\to\infty}\left(\sum_{j=0}^n \hat{\boldsymbol{x}}(j)\mathscr{L}_j\right)^i\right) \qquad \text{in } L_2([0,1]),$$

with $\|x^i\| = \|\hat{\boldsymbol{x}}_i\|$ and $\hat{\boldsymbol{x}}_i = \boldsymbol{\Delta\gamma}_i$. In other words, $\hat{\boldsymbol{x}}^{(i)} = \hat{\boldsymbol{x}}_i$ or equivalently, $\hat{\boldsymbol{x}}^{(i)} = \boldsymbol{\Delta\gamma}_i = (\boldsymbol{\Delta\gamma}_1)^{(i)}$, which is (3.4).

We next prove the (b) part. By the disintegration (3.1) of the measure μ,

$$\gamma_i(j) = \int_0^1 t^j f_i(t)\,dt, \qquad i,j=0,1,\ldots$$

for some nonnegative measurable functions $f_i\in L_\infty([0,1])$, $i=1,2,\ldots$.

As (3.5) holds one may conclude that $\hat{\boldsymbol{q}}_i = \hat{\boldsymbol{q}}_1^{(i)}$ where $\hat{\boldsymbol{q}}_i$ is the (shifted) Legendre–Fourier vector of coefficients associated with $f_i\in L_2([0,1])$, $i=1,\ldots$. Hence by Proposition 2.2, $f_i(t) = f_1(t)^i$ a.e. on $[0,1]$, for every $i=1,2,\ldots$. That is, for every $i=1,2,\ldots$, there exists a Borel set $B_i\subset[0,1]$ with Lebesgue measure zero such that $f_i(t)=f_1(t)^i$ for all $t\in[0,1]\backslash B_i$. Therefore the Borel set $B=\cup_{i=1}^\infty B_i$ has Lebesgue measure zero and for all $i=1,2,\ldots$,

$$f_i(t) = f_1(t)^i, \qquad \forall\, t\in[0,1]\backslash B.$$

Hence for every $t\in[0,1]\backslash B$,

$$\int_{[0,1]} x^i\psi(dx|t) = f_1(t)^i = \int_{[0,1]} x^i\, d\delta_{f_1(t)}, \qquad \forall\, i=1,2,\ldots.$$

where $\delta_{f_1(t)}$ is the Dirac measure at the point $f_1(t)\in[0,1]$. As measures on compact sets are moment determinate, one must have $\psi(dx|t)=\delta_{f_1(t)}$, for all $t\in[0,1]\backslash B$. Therefore $d\mu(x,t) = \delta_{f_1(t)}\,dt$, i.e., the measure μ is supported on the trajectory $\{(t,x(t))\colon t\in[0,1]\}$, where $x(t) = f_1(t)$ for almost all $t\in[0,1]$. ∎

Remark 3.2. If the trajectory $t \mapsto x(t)$ is a polynomial of degree say d, then the vector of Legendre–Fourier coefficients $\hat{\boldsymbol{x}} \in \ell^2$ has at most $d+1$ non-zero elements. Therefore for every $j = 2, \ldots, \hat{\boldsymbol{x}}_j \in \ell^2$ also has at most $jd+1$ non-zero elements and the condition (3.5) can be checked easily.

In Theorem 3.1(a) one assume that $\sup_n \left\| \sum_{j=0}^{n} \hat{\boldsymbol{x}}(j) \mathscr{L}_j \right\|_\infty < \infty$ which is much weaker than, e.g., assuming the uniform convergence $\left\| \sum_{j=0}^{n} \hat{\boldsymbol{x}}(j) \mathscr{L}_j - x \right\|_\infty \to 0$ as $n \to \infty$. The latter (which is also much stronger than the a.e. pointwise convergence) can be obtained if the function $x(t)$ has some smoothness properties. For instance if x belongs to some Lipschitz class of order larger then or equal to $1/2$, then uniform convergence takes place and one may even obtain rates of convergence; see, e.g., [12] and also [15] for a comparison (in terms of convergence) of Legendre and Chebyshev expansions. In fact, quoting the authors of [5], *"... knowledge of the partial spectral sum of an L_2 function in $[-1,1]$ furnishes enough information such that an exponential convergent approximation can be constructed in any subinterval in which f is analytic".*

Example 3.3. To illustrate Theorem 3.1 consider the following toy example with μ on $[0,1]^2$ and with marginal w.r.t. "t" being the uniform distribution on $[0,1]$ and conditional $\psi(dx|t) = \delta_{\exp(-t)}$ for all $t \in [0,1]$. That is, $t \mapsto x(t) = \exp(-t)$.

Then the first 11 Legendre–Fourier coefficients $\hat{\boldsymbol{x}}(j)$, $j = 0, \ldots, 10$ of x read

$$\begin{aligned} \hat{\boldsymbol{x}} = [&0.63212055 \quad -0.1795068 \quad 0.0230105 \quad -0.0019370 \quad 0.0001217 \\ &-0.0000061 \quad 0.0000002 \quad -0.00000001 \quad -0.00000004 \quad -0.0000015 \quad -0.0000625]. \end{aligned}$$

Similarly the first 11 Legendre–Fourier coefficients of $t \mapsto x(t)^2 = \exp(-2t)$ read

$$\begin{aligned} \hat{\boldsymbol{x}}_2 = [&0.4323323 \quad -0.2344075 \quad 0.0588678 \quad -0.0097965 \quad 0.0012219 \\ &-0.0001219 \quad 0.0000101 \quad -0.0000007 \quad 0.00000004 \quad -0.000000004 \quad -0.00000004]. \end{aligned}$$

With $n = 5$ the polynomial $t \mapsto x_5^{(2)}(t) := \left(\sum_{k=0}^{5} \hat{\boldsymbol{x}}(j) \mathscr{L}_j(t) \right)^2$ reads

$$\begin{aligned} x_5^{(2)}(t) &= \sum_{k=0}^{10} \hat{\boldsymbol{x}}_5^{(2)}(j)\, \mathscr{L}_j(t), \qquad t \in \mathbb{R}, \qquad \text{with} \\ \hat{\boldsymbol{x}}_5^{(2)} &= [0.4323336 \quad -0.2344129 \quad 0.0588626 \quad -0.0097976 \quad 0.0012219 \\ &\quad -0.0001218 \quad 0.0000098 \quad -0.0000006 \quad 0.00000003 \quad -0.000000001 \quad 0.0000000], \end{aligned}$$

and we can observe that $\hat{\boldsymbol{x}}_5^{(2)} - \hat{\boldsymbol{x}}_2 \approx O(10^{-5})$.

In fact the curves $t \mapsto x_5^{(2)}(t)$ and $t \mapsto \exp(-2t)$ are almost indistinguishable on the interval $[0,1]$.

3.1 A more general case

We have considered a measure μ on $[0,1]^2$ whose marginal with respect to $t \in [0,1]$ is the Lebesgue measure. The conditions of Theorem 3.1 are naturally stated in terms of the (shifted) Legendre–Fourier coefficients associated with the functions $t \mapsto f_i(t)$ of $L_2([0,1])$ defined in (3.1).

However, the same conclusions also hold if the marginal of μ with respect to $t \in [0,1]$ is some measure $d\nu = h(t)dt$ for some nonnegative function $h \in L_1([0,1])$ with all moments finite. The

only change is that now we have to consider the orthonormal polynomials $t \mapsto \mathcal{H}_j(t)$, $j = 0, 1, \ldots$, with respect to ν. Recall that all the $\mathcal{H}_j$'s can be computed from the moments

$$\gamma_0(j) = \int t^j d\mu(x,t) = \int_0^1 t^j \, d\nu(t) = \int_0^1 t^j h(t) \, dt, \qquad j = 0, 1, \ldots .$$

Then proceeding as before, for every $i = 1, 2, \ldots$,

$$\int t^j x^i \, d\mu(x,t) = \int_0^1 t^j \underbrace{\left(\int_{[0,1]} x^i \psi(dx|t) \right)}_{= f_i(t) \in L_2([0,1],\nu)} h(t) \, dt, \qquad j = 0, 1, \ldots ,$$

and we now consider the vector of coefficients $\hat{\boldsymbol{f}}_{hi} = (\hat{\boldsymbol{f}}_{hi}(j))$ defined by

$$\hat{\boldsymbol{f}}_{hi}(j) = \int_0^1 \mathcal{H}_j(t) f_i(t) h(t) \, dt, \qquad j = 0, 1, \ldots$$

the analogues for the measure $d\nu = h(t)dt$ and the function $f_i \in L_2([0,1],\nu)$, of the (shifted) Legendre–Fourier coefficients $\hat{\boldsymbol{f}}_{\boldsymbol{i}}(j)$ of f_i in (3.1) for the Lebesgue measure on $[0,1]$. Then the conditions in Theorem 3.1 would be exactly the same as before, excepted that now, $\hat{\boldsymbol{x}} = (\hat{\boldsymbol{x}}(j))$ with

$$\hat{\boldsymbol{x}}(j) = \int_0^1 \mathcal{H}_j(t) x(t) h(t) \, dt, \qquad j = 0, 1, \ldots .$$

3.2 Discussion

Theorem 3.1 may have some practical implications. For instance consider the *weak* formulation $\mathcal{P}$ of an optimal control problem $\mathbf{P}$ as an infinite-dimensional optimization problem on an appropriate space of (occupation) measures, as described, e.g., in [14]. In [8] the authors propose to solve a hierarchy of semidefinite relaxations $(\mathcal{P}_k)$, $k = 1, 2, \ldots$, of $\mathcal{P}$. Each optimal solution of $\mathcal{P}_k$ provides with a finite sequence $\mathbf{z}^k = (z^k_{j,\alpha,\beta})$ such that when $k \to \infty$, $\mathbf{z}^k \to \mathbf{z}^*$ where $\mathbf{z}^*$ is the infinite sequence of some measure $d\mu(t, \mathbf{x}, \mathbf{u})$ on $[0,1] \times \mathbf{X} \times \mathbf{U}$, where $\mathbf{X} \subset \mathbb{R}^n$, $\mathbf{U} \subset \mathbb{R}^m$, are compact sets.

Under some conditions both problems $\mathbf{P}$ and its relaxation $\mathcal{P}$ have same optimal value. If μ is supported on feasible trajectories $\{(t, \mathbf{x}(t), \mathbf{u}(t)) \colon t \in [0,1]\}$ then these trajectories are optimal for the initial optimal control problem $\mathbf{P}$. So it is highly desirable to check whether indeed μ is supported on trajectories from the only knowledge of its moments $\mathbf{z}^* = (z^*_{j,\alpha,\beta})$.

By construction of the moment sequences $\mathbf{z}^k$ one already knows that the marginal of μ with respect to the variable "t" is the Lebesgue measure on $[0,1]$. Therefore we are typically in the situation described in the present paper. Indeed to check whether μ is supported on trajectories $\{(t, (x_1(t), \ldots, x_n(t), u_1(t), \ldots, u_m(t)) \colon t \in [0,1]\}$, one considers each coordinate $x_i(t)$ or $u_j(t)$ separately. For instance, for $x_i(t)$ one considers the subset of moments $\gamma_k(j) = (z^*_{j,\alpha,0})$ with $j = 0, 1, \ldots$, $\alpha = (0, \ldots, 0, k, 0, \ldots, 0) \in \mathbb{N}^n$, $k = 0, 1, \ldots$, with k in position i. If (3.4) holds then indeed the marginal μ_{t,x_i} of μ on (t, x_i), with moments $(\gamma_k(j))$ is supported on a trajectory $\{(t, x_i(t)) \colon t \in [0,1]\}$.

Of course, in (3.4) there are countably many conditions to check whereas in principle only finitely many moments of z^* are available (and with some inaccuracy due to (a) solving numerically a truncation $\mathcal{P}_k$ of $\mathcal{P}$, and (b) the convergence $\mathbf{z}^k \to \mathbf{z}^*$ has not taken place yet). So an issue of future investigation is to provide necessary (or sufficient?) conditions based only on finitely many (approximate) moments of μ.

Acknowledgements

Research funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement ERC-ADG 666981 TAMING).

References

[1] Beiglböck M., Griessler C., A land of monotone plenty, Technical report, Department of Mathematics, University of Vienna, 2015.

[2] Charina M., Lasserre J.B., Putinar M., Stöckler J., Structured function systems and applications, *Oberwolfach Rep.* **10** (2013), 579–655.

[3] Diaconis P., Freedman D., The Markov moment problem and de Finetti's theorem. I, *Math. Z.* **247** (2004), 183–199.

[4] Dunkl C.F., Xu Y., Orthogonal polynomials of several variables, *Encyclopedia of Mathematics and its Applications*, Vol. 81, Cambridge University Press, Cambridge, 2001.

[5] Gottlieb D., Shu C.-W., On the Gibbs phenomenon. III. Recovering exponential accuracy in a sub-interval from a spectral partial sum of a piecewise analytic function, *SIAM J. Numer. Anal.* **33** (1996), 280–290.

[6] Helton J.W., Lasserre J.B., Putinar M., Measures with zeros in the inverse of their moment matrix, *Ann. Probab.* **36** (2008), 1453–1471, math.PR/0702314.

[7] Lasserre J.B., Borel measures with a density on a compact semi-algebraic set, *Arch. Math. (Basel)* **101** (2013), 361–371, arXiv:1304.1716.

[8] Lasserre J.B., Henrion D., Prieur C., Trélat E., Nonlinear optimal control via occupation measures and LMI-relaxations, *SIAM J. Control Optim.* **47** (2008), 1643–1666, math.OC/0703377.

[9] McCann R.J., Guillen N., Five lectures on optimal transportation: geometry, regularity and applications, arXiv:1011.2911.

[10] Putinar M., Extremal solutions of the two-dimensional L-problem of moments, *J. Funct. Anal.* **136** (1996), 331–364.

[11] Putinar M., Extremal solutions of the two-dimensional L-problem of moments. II, *J. Approx. Theory* **92** (1998), 38–58, math.CA/9512222.

[12] Suetin P.K., On the representation of continuous and differentiable functions by Fourier series in Legendre polynomials, *Sov. Math. Dokl.* **158** (1964), 1408–1410.

[13] Villani C., Topics in optimal transportation, *Graduate Studies in Mathematics*, Vol. 58, Amer. Math. Soc., Providence, RI, 2003.

[14] Vinter R., Convex duality and nonlinear optimal control, *SIAM J. Control Optim.* **31** (1993), 518–538.

[15] Wang H., Xiang S., On the convergence rates of Legendre approximation, *Math. Comp.* **81** (2012), 861–877.